New Phytologist

From the Arctic to the tropics: multi-biome prediction of leaf mass per area using leaf optical properties

Journal:	*New Phytologist*
Manuscript ID	NPH-MS-2019-29464
Manuscript Type:	MS - Regular Manuscript
Date Submitted by the Author:	n/a
Key Words:	spectroscopy, remote sensing, plant traits, leaf mass area, specific leaf area, partial least squares regression (PLSR)

Complete List of Authors:
Serbin, Shawn; Brookhaven National Laboratory, Environmental & Climate Sciences Department; Stony Brook University, Ecology and Evolution
Wu, Jin; Brookhaven National Laboratory, Environmental and Climate Sciences
Ely, Kim; Brookhaven National Laboratory, Environmental & Climate Sciences Department
Kruger, Eric; University of Wisconsin-Madison, Forest and Wildlife Ecology
Townsend, Philip; University of Wisconsin - Madison, Forest and Wildlife Ecology
Meng, Ran; Brookhaven National Laboratory, Department of Environmental & Climate Sciences
Wolfe, Brett; Smithsonian Tropical Research Institute, Botany
Chlus, Adam; University of Wisconsin - Madison, Forest and Wildlife Ecology
Wang, Zhihui; University of Wisconsin - Madison, Forest and Wildlife Ecology
Rogers, Alistair; Brookhaven National Laboratory, Environmental & Climate Sciences Department

From the Arctic to the tropics: multi-biome prediction of leaf mass per area using leaf optical properties

Shawn P. Serbin[1*], Jin Wu[1], Kim S. Ely[1], Eric L. Kruger[2], Philip A. Townsend[2], Ran Meng[1], Brett T. Wolfe[3], Adam Chlus[2], Zhihui Wang[2], Alistair Rogers[1]

[1]Environmental and Climate Sciences Department, Brookhaven National Laboratory, Upton, NY 11973, U.S.A.
[2]Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, Madison, WI 53706, U.S.A.
[3]Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Panama

Author	Orchid Identifier
Shawn Serbin	0000-0003-4136-8971
Jin Wu	0000-0001-8991-3970
Kim Ely	0000-0002-3915-001X
Eric Kruger	no orchid identifier
Philip Townsend	0000-0001-7003-8774
Ran Meng	0000-0003-4756-9934
Brett Wolfe	0000-0001-7535-045X
Adam Chlus	no orchid Identifier
Zhihui Wang	0000-0003-1064-7820
Alistair Rogers	0000-0001-9262-7430

Corresponding Author
Shawn Paul Serbin
Environmental and Climate Sciences Department
Brookhaven National Laboratory
Upton, NY 11973-5000
USA
Email: sserbin@bnl.gov
Phone: +1.631.344.3165

Keywords: spectroscopy, remote sensing, plant traits, leaf mass area, specific leaf area, partial least squares regression (PLSR)

Type of Paper: Original Research

Word Count:

Summary: 199 words

Main-Text--

Introduction: 1131 words

Materials and Methods: 1795 words

Results and Discussion: 2968 words

Acknowledgement: 130 words

4 Figures

Supplementary materials: 11 Figures 1 Table

Summary

Leaf-mass-per-area (LMA) is a key plant trait, reflecting tradeoffs between leaf photosynthetic function, longevity, and structural investment. Capturing spatial and temporal variability in LMA has been a long-standing goal of ecological research and is an essential component for advancing Earth system models. Despite the substantial variation in LMA within and across Earth's biomes, an efficient, globally generalizable approach to predict LMA is still lacking.

We explored the capacity to predict LMA from leaf spectra across much of the global LMA trait space, with values ranging from 17 g m^{-2} to 393 g m^{-2}. Our dataset contained leaves from a wide range of biomes from the high Arctic to the tropics, included broad- and needle-leaf species, and upper and lower canopy (i.e. sun and shade) growth environments.

Here we demonstrate the capacity to rapidly estimate LMA using only spectral measurements across a wide range of species, leaf age, and canopy position, from diverse biomes. Our model captures LMA variability with high accuracy and low error (R^2=0.89; RMSE=15.45 g m^{-2})

Our finding highlights that the leaf economic spectrum is mirrored by the leaf optical spectrum paving the way for this technology to predict the functional diversity of ecosystems across global biomes.

Introduction

A key objective of plant ecology is to characterize the functional diversity of plants that have evolved different strategies for growth, reproduction and for coping with biotic and abiotic stressors (Wright *et al.*, 2004). Accurate characterization of this functional diversity in Earth system models (ESMs) will improve our ability to model the Earth system and understand the effect of global change on the cycling and storage of carbon, water and energy (Pavlick *et al.*, 2013; Fisher *et al.*, 2018). Therefore, an increasing number of ESMs are moving toward incorporating approaches that require a broader and more comprehensive representation of plant trait variation within and across biomes (Xu *et al.*, 2012; van Bodegom *et al.*, 2014; Wullschleger *et al.*, 2014; Fisher *et al.*, 2015; Fisher *et al.*, 2018). As a result, considerable effort has been invested in the measurement and monitoring of plant traits across a range of biomes, and the storage and synthesis of that information in global databases (Wright *et al.*, 2004; Kattge *et al.*, 2011; Lebauer *et al.*, 2013; Butler *et al.*, 2017). Yet, the high degree of plant functional diversity and plasticity makes this apparently simple goal extremely challenging (Reich *et al.*, 1997; Reich *et al.*, 1999; Serbin *et al.*, 2014; Wu *et al.*, 2017; Osnas *et al.*, 2018). Consequently, the extent of global trait coverage is still woefully inadequate (Schimel *et al.*, 2015).

In recent decades, remote sensing has shown increasing promise as a means to capture plant traits across scales using spectroscopic approaches. For example, several recent studies have highlighted the capacity to connect remotely sensed spectra to characterize variation in a number of key functional traits across individual leaves, canopies, and landscapes (Dahlin *et al.*, 2013; Asner *et al.*, 2015; Singh *et al.*, 2015; Shiklomanov *et al.*, 2016; Yang *et al.*, 2016; Wu *et al.*, 2017). Approaches have included empirical spectra-trait modeling, such as partial least-squares regression (e.g. Serbin *et al.*, 2014; Singh *et al.*, 2015; Yendrek *et al.*, 2017) and spectral vegetation indices (SVIs, e.g. Feret *et al.*, 2011), as well as semi-mechanistic approaches, including the use of the PROSPECT leaf-level radiative transfer model (RTM, e.g. Shiklomanov

116 *et al.*, 2016; Féret *et al.*, 2017). In general, all of these approaches rely on the fundamental
117 biophysical connection between leaf chemistry and structure and the resultant optical properties
118 of plants (Curran, 1989; Ustin *et al.*, 2004; Kokaly *et al.*, 2009; Ollinger, 2011), and as a result,
119 could be used to fill critical gaps in our understanding of the variation in plant traits over
120 landscapes and biomes (Schimel *et al.*, 2015; Jetz *et al.*, 2016).

121 However, most studies that have illustrated the strong promise of remote sensing to
122 estimate foliar traits across scales have focused on a relatively narrow subsample of geographic
123 regions (e.g. Townsend *et al.*, 2003; Feret *et al.*, 2011; Dahlin *et al.*, 2013; Asner *et al.*, 2015;
124 Yang *et al.*, 2016), capturing only a small fraction of total trait and spectral space. In addition,
125 semi-mechanistic methods such as PROSPECT may have other assumptions or limitations that
126 could inhibit their broad application, for example, a limited number of supported traits (Féret *et al.*,
127 2017) or the generally poorer performance for needle-leaf species (Malenovský *et al.*, 2006;
128 Shiklomanov *et al.*, 2016). As a result, the general applicability of existing spectra-trait models
129 and approaches across biomes and in the wider trait space is not known. In many cases those
130 models trained across a limited trait space have been shown to break down when applied more
131 broadly (e.g. Sims & Gamon, 2002; Gitelson *et al.*, 2003; le Maire *et al.*, 2008).

132 Among plant traits, foliar morphology is commonly characterized using leaf mass per area
133 (LMA)—the ratio of a leaf's dry mass to its surface area (g dry mass per m^2 leaf area)—or its
134 reciprocal, specific leaf area (SLA). LMA captures the trade-off between a plant's investment in
135 leaf structure and turnover rate versus leaf surface area and light harvesting for photosynthesis
136 (Wright *et al.*, 2004; Shipley *et al.*, 2006; Poorter *et al.*, 2009). Given its strong linkage with plant
137 function (Reich *et al.*, 1997; Wright *et al.*, 2004; Serbin *et al.*, 2012; Osnas *et al.*, 2018), LMA is a
138 critical parameter in plant ecology. Illustrations of the importance of LMA include its use as a basis
139 for monitoring biodiversity (Skidmore *et al.*, 2015), its role in modeling canopy radiation transfer
140 (Jacquemoud *et al.*, 2009; Ollinger, 2011), and its widespread use as an input in ecosystem
141 process models (Fisher *et al.*, 2014; Xu *et al.*, 2016; Ricciuto *et al.*, 2018)

Importantly, the substantial global variation in LMA, which ranges from 14 to 1500 g m^{-2} globally (Wright *et al.*, 2004), exists within and across species (Castro-Díez *et al.*, 2000; Wright *et al.*, 2004; Paula & Pausas, 2006; Poorter *et al.*, 2009) and is mediated by local gradients in light, water and nutrient availability (Niinemets, 2007; de la Riva *et al.*, 2016; Liu *et al.*, 2017), leaf age (Wu *et al.*, 2016), as well as acclimation and adaptation to short- and long-term climate dynamics (Volin *et al.*, 2002; Paula & Pausas, 2006; Poorter *et al.*, 2009). LMA variation is attributed primarily to differences in leaf density and volume-to-area ratio (Poorter *et al.*, 2009; John *et al.*, 2017). The fundamental information on these attributes is found in the reflectance spectrum of a leaf, which captures its physical properties—e.g., thickness, density, the depth of palisade layers, albedo and elemental composition (Ollinger, 2011). Given the co-ordination between leaf traits and optical properties, we hypothesize that, using a spectroscopy approach, we can collapse the vast structural and functional diversity of leaves from different plant species, leaf types (e.g. grasses, forbs, broadleaf and needle-leaf evergreen trees), across a wide range of environments, into a single generalizable model.

Here we present a broad, multi-biome analysis linking fundamental co-variation in LMA and spectroscopic (also known as hyperspectral) reflectance. We use a large dataset to develop a robust statistical model to infer LMA from corresponding leaf optical properties, and then validate this cross-biome model using independent datasets of additional spectra and LMA observations from a similar range of plant material as well as from external validation sources. The core training and validation datasets include leaves from the high Arctic to the tropics, and contains measurements from grasses, forbs, deciduous and evergreen shrubs, deciduous and evergreen broad leaved trees, needle leaved trees, and crop species. It spans a highly diverse range of 1) leaf morphologies, including glabrous, highly reflective and waxy leaf types; 2) microenvironment, including measurements from upper-canopy, sunlit leaves, and lower-canopy, shaded leaves; 3) developmental stage, including recently emerged, mature and old leaves; and

4) elevation, including measurements from sea level to more than 2,000 m. Thus, our analysis is based on data representing a large fraction of the global trait space for LMA.

Materials and Methods

Plant Material

We collected and assembled a large dataset (n = 2,478 leaves from more than 176 species) of combined LMA and leaf reflectance spectra for model development from the high Arctic in northern Alaska to the tropics in central America and Brazil (Figure 1). Our sites are distributed across a large proportion of the Earth's habitable climate space, encompassing a ~40°C range of mean annual temperatures and a ~200 – 2400 mm/year range in mean annual precipitation. Our sites in Alaska include coastal tundra vegetation within the Barrow Environmental Observatory (BEO), near Barrow (now Utqiaġvik) Alaska (Brown *et al.*, 1980; Rogers *et al.*, 2017b) and dwarf and tall shrub vegetation on the Seward Peninsula (Rogers *et al.*, 2016; Serbin & Rogers, 2019). Our sites located in the upper midwest and northeastern US are dominated by northern temperate forest species, including deciduous broadleaf hardwoods and evergreen, needle-leaf conifers (Serbin *et al.*, 2014). Measurements on Mediterranean and agricultural plants were conducted in the Coachella and central valleys of California (Serbin *et al.*, 2015), and across an elevation gradient of sites ranging from woodlands to alpine forests in the Sierra Mountains (Goulden *et al.*, 2012; Dubois *et al.*, 2018). The data for tropical species were collected in several separate locations; a seasonal, wet evergreen Amazonian forest near Santarem, Brazil (Wu *et al.*, 2017), a seasonal, wet evergreen forest in the San Lorenzo Protected Area and a seasonal, dry forest in the Parque Natural Metropolitano near Panama City in The Republic of Panama (Wright *et al.*, 2003) and a collection of plants grown in an artificial tropical forest within the Biosphere 2 facility (Walter & Carmen Lambrecht, 2004). All the data used in this study are publicly available, and the published data sets include detailed descriptions of the sampling locations (Table S1).

193 *Spectroscopic measurements*

194 The approach, instrumentation and personnel used to measure leaf spectral reflectance
195 varied across the different sites and projects included in this study. The leaf reflectance of Arctic
196 plants was measured using an Analytical Spectral Devices (ASD Inc., Longmont Colorado, USA)
197 FieldSpec 3 with the ASD leaf clip assembly, or a Spectra Vista Corporation (Poughkeepsie, NY,
198 USA) HR-1024i full spectrum spectroradiometer together with the standard leaf clip assembly as
199 well as the more recent LC-RP-Pro leaf clip. Leaf spectra collected in the Upper Midwest and
200 northeastern USA were measured with an ASD FieldSpec pro or FieldSpec 3 with the ASD leaf
201 clip assembly. The data collected in California was measured using either an ASD FieldSpec 3
202 or a Spectral Evolution (SE, Lawrence, Massachusetts) PSR-3500 full range spectrometer with
203 an attached SE leaf clip assembly. Data from Biosphere2 and Brazil were measured with an ASD
204 FieldSpec Pro and ASD leaf clip, the data from Puerto Rico with a SE PSR+ with the SVC LC-
205 RP-Pro and a custom fiber optic, while data from The Republic of Panama were measured with
206 an SVC HR-1024i with the LC-RP-Pro attachment. All instruments had leaf clip assemblies that
207 contained an internal calibrated light source, and all reflectance measurements were referenced
208 against a 99% Spectralon reflectance standard (type varied across sites). For needle-leaf, or very
209 small leaf species, we used needle or leaf mats (Serbin, 2012). Needles of a similar age (i.e.
210 current year, previous year, and older) were laid out edge to edge creating a single layer and
211 taped at the ends to hold the needles tightly together before inserting the mat into the leaf
212 clip. No tape was visible to the fiber optic inside the leaf clip. All spectral measurements were
213 processed using the R-FieldSpectra package (https://github.com/serbinsh/R-FieldSpectra). For
214 SVC data, however, we first corrected the discontinuities in the spectra in the detector overlap
215 areas using the vendor provided software. Measurements at the edges of the spectral range of
216 these spectrometers (350-500 nm, > 2400 nm) suffer from low signal-to-noise thus were excluded
217 from the analysis.
218

Measurements of LMA

Measurements of LMA also varied among the different studies and biomes. Various methods were used to determine leaf area, including measurement of length and width with a ruler and a hand lens (graminoid species), leaf disc punches (broadleaf species) and optical approaches. For small or compound leaves where it was not possible to take discs, area was measured with a leaf area meter (LI-3100C Area Meter, LI-COR, Lincoln, NE) calibrated for use with high edge-to-area ratio leaves and operated at high (0.1 mm^2) resolution, or a flatbed scanner (evergreen needle-leaf trees) followed by area estimation using ImageJ (Serbin *et al.*, 2014). Leaves and leaf sections of known area were then dried to constant mass in a ventilated oven (60°C to 70°C) and leaf dry mass measured on a top-pan balance. All leaf area data was provided on a projected leaf area basis.

LMA partial least-squares regression modeling

To relate the variability in LMA across sites, species, and environments we utilized a Partial Least-Squares Regression (PLSR) modeling approach (Geladi & Kowalski, 1986; Wold *et al.*, 2001) using the PLS package (Mevik & Wehrens, 2007) in the R open source statistical environment (R Core Team, 2017). PLSR is widely utilized in spectroscopy and chemometric analyses given its ability to handle high predictor collinearity as well as a large number of predictor variables that may exceed the number of observations. PLSR minimizes the implications of these circumstances by reducing the number of predictor variables down to a relatively few, orthogonal latent components (Geladi & Kowalski, 1986; Wold *et al.*, 2001). Moreover, PLSR does not assume the measurement of predictor variables (reflectance values at given wavelengths in this case) was made without error.

Our PLSR model development has been described previously (Serbin *et al.*, 2014) and is briefly summarized here. We first applied a square root transformation to the LMA data to reduce the right skewness distribution of the original data (Figure S1) and satisfy the required normal

distribution for the PLSR analysis. We then randomly split the full dataset into calibration (80%, n=1978) and independent validation (20%, n=500) subsets (Figure S2), ensuring that both subsets spanned the full range of observations and included measurements from each study and sample location. Using just the calibration dataset we developed the multi-biome leaf spectra-LMA model and we then tested this final model using the validation data. To avoid the potential to over-fit the spectra-LMA calibration model, we optimized the number of PLSR components in the model by minimizing the prediction residual sum of squares (PRESS) statistic (Chen *et al.*, 2004). We calculated the PRESS statistic of successive model components through a cross-validation analysis where we minimized the PRESS statistic until successive PLSR components did not reduce model predictive error as assessed using a *t* test (Serbin et al., 2014). Lastly, we calculated the variable influence on projection (VIP) metric (Wold *et al.*, 2001) of the final model to identify the regions of the spectrum that contributed significantly to the prediction of LMA (Serbin *et al.*, 2014).

In addition to the general model development, we also conducted a PLSR model uncertainty analysis to characterize the predictive uncertainty, given the variability and error in measured LMA, spectra, and the relationship between the two. This was done by splitting the original calibration dataset into a balanced 70% to 30% via 1000 permutations and generating the same number of new model coefficient vectors, following Serbin et al. (2014). The result of this uncertainty analysis is an ensemble of PLSR models that can be used to predict new values of LMA based on spectral measurements plus the predictive uncertainty for each new value.

Finally, we quantified the performance of the multi-biome LMA PLSR model using the independent validation dataset. In this step, we validated the model and examined the residuals for model bias and predictive performance. We used four main evaluation metrics: the coefficient of determination (R2), RMSE, and the residual and regression biases. All model and error results presented below are shown in original LMA units not the square root transformed units that are the initial output of the PLSR model.

Examples of applying the PLSR model to new observations

We provide a simple R script (Supplemental R script) to illustrate the utility of our model and how it can be used to estimate LMA values from leaf spectral reflectance observations not used for model development. This script can be run to automatically download the foundational LOPEX (Leaf Optical Properties Experiment) and ANGERS spectral datasets (Hosgood *et al.*, 1994; Jacquemoud *et al.*, 2003), apply the PLSR model, and provide the results. The LMA model coefficients are provided through GitHub (https://tinyurl.com/y8pek6n3) and the leaf spectral data are provided from the Ecological Spectral Information System (EcoSIS) database (https://ecosis.org).

Furthermore, we also provide an extensive, external validation of our multi-biome PLSR model with data collected in the upper Midwest, U.S. and nine NEON (National Ecological Observatory Network) locations spanning seven NEON domains. The first data set consisted of fully expanded, peak greenness samples collected in and around the Madison, Wisconsin area and the University of Notre Dame Environmental Research Center (UNDERC) between June and September 2017 (Chlus et al., unpublished data). A mix of sunlit and shaded foliar samples were collected from broadleaf trees (n=7446), graminoids (n=74), forbs (n=2017), and vine (n=218) species across the growing season. All samples were immediately scanned using an ASD FieldSpec 3, Spectra Vista (SVC) 1024i and Spectral Evolution PSR+ with their respective leaf contact probes and external light sources. A Spectralon white reference was scanned before each sample to calculate relative reflectance. After spectral measurement, leaf area of each sample was immediately recorded using the LI-3100 benchtop leaf scanner (Li-Cor Biosciences, Lincoln, NE). Samples were subsequently flash-frozen in liquid nitrogen, freeze-dried in a VirTis lyophilizer (SP Scientific, Gardiner, NY), and weighed with a precision balance.

The second data set consisted of peak greenness foliar samples collected during the summer of 2017 at NEON (National Ecological Observatory Network) sites in Wisconsin, Alabama, Georgia, Florida, Virginia, Maryland, Tennessee, Kansas and North Dakota, and also

297 included trees (n=2584), graminoids (n=381) and forbs (195) species (Wang et al., unpublished
298 data). Spectra were collected immediately using the ASD FieldSpec and/or Spectral Evolution
299 PSR+ with their leaf contact probes. Leaves were scanned on a 600 dpi flatbed scanner (Epson,
300 Nagano, Japan), oven-dried at 65 °C for 48 h to a constant mass and weighed on the precision
301 balance. The only processing applied to the spectra was the removal of spectral discontinuities
302 (following Serbin *et al.* (2012)) at approximately 1000 and 1900 nm where "jumps" sometimes
303 occur at overlapping wavelengths between detectors within the instruments. The final multi-biome
304 PLS equation was applied to all spectra and compared with laboratory measurements.
305
306 *Data Availability*
307 All data used in this manuscript are publicly available through online data portals, including the
308 U.S. Department of Energy (DOE) NGEE-Arctic and NGEE-Tropics data portals as well as the
309 EcoSIS spectral database (Supplemental Table 1).
310
311 **Results and Discussion**
312 We employed an extensive dataset of leaf spectral reflectance and LMA across multiple
313 biomes and spanning a large range of habitable climate space (Figure 1) to develop a generalized
314 approach to model the variation in LMA (Figure 2) using only leaf optical properties (Figure 3).
315 We found a very strong capacity for the empirical PLSR spectra-trait modeling approach to
316 accurately model multi-biome variation in LMA using spectral reflectance (Figure 4). Our results
317 show that spectra alone can explain 89% of the variation in LMA with a low model bias (0.96 g
318 m^{-2}) and RMSE (15.45 g m^{-2}) when compared with our core validation dataset. We provide more
319 details on our input datasets, PLSR modeling, and validation below as well as a discussion of our
320 work in the larger functional trait and remote sensing research communities.
321
322

Biotic and abiotic variation in LMA

Across our sample sites (Figure 1), we observed a broad range of LMA in our model development dataset, with values ranging from 17 g m^{-2} to 393 g m^{-2} (Figure 2, Figure S1). As expected, mean top-of-canopy LMA of the needle-leaf evergreen conifer species (mean = 165 g m^{-2} ± 62 S.D.), was significantly larger than broadleaf deciduous species, including grasses and forbs (67 g m^{-2} ± 24 S.D.). The LMAs of evergreen broadleaf species from arid environments (169 g m^{-2} ± 72 S.D.) were similar to those of needle-leaf evergreen species, consistent with past observations (Paula & Pausas, 2006; de la Riva *et al.*, 2016), and owing to the well-documented differences in leaf lifespans that typically require higher resource investment in leaf construction (Wright *et al.*, 2004; Poorter *et al.*, 2009). For the tropical species, LMA averaged 106 g m^{-2} (± 37 S.D.) and ranged from 22 to 306 g m^{-2}.

Within a biome, LMA variation (Figure S3) was large and related to biotic differences across species as well as factors such as leaf and plant age and abiotic factors such as the leaf growth environment, including position in the canopy (Serbin *et al.*, 2014; Wu *et al.*, 2017), consistent with other studies (Niinemets *et al.*, 2015; John *et al.*, 2017; Osnas *et al.*, 2018). Notably, variation within species was often as large as that across species (data not shown), as has been noted in previous work (Butler *et al.*, 2017). On average, LMA in upper-canopy, fully sunlit leaves were 36% (± 20 S.D.) higher than in fully shaded leaves of the same species, but this difference varied with growth form and leaf habit, likely owing to known differences in growth strategies, seasonality and the difference in the light penetration through a broadleaf or needle-leaf canopy (Wright *et al.*, 2004; Ollinger, 2011; Butler *et al.*, 2017). Overall, the range in LMA within and across biomes showed similarities and overlap from the Arctic to the tropics, despite the vastly different leaf morphologies resulting from environmental and biological drivers, and these similarities were likely related to the underlying variation in leaf density and volume to area ratios, which can lead to similar values of LMA with significantly different leaf geometries (Poorter *et al.*, 2009; John *et al.*, 2017).

Variation in leaf reflectance across species, biomes, and growth environment

As with LMA, we observed significant variation in the measured leaf-level spectral reflectance in our model development dataset within and across biomes (Figure 3ab; Figure S4). The biome-level spectral reflectance displayed a similar shape and comparable magnitude across the spectral range examined here (500-2400 nm, Figure S4), however the within-biome variation (presented as the coefficient of variation in reflectance by wavelength) showed some key differences (Figure 3b) in the visible (i.e. 500-700 nm) and shortwave infrared (SWIR) regions (e.g. 1900-2500 nm). Similarity the biome-level mean reflectance appeared similar due to the large variation in reflectance across wavelengths within each biome (Figure S4), which mirrored the considerable within-biome variation in LMA (Figure 2; Figure S3). This pattern has been observed in past research for similar vegetation types (Asner & Martin, 2008; Feret *et al.*, 2011; Yang *et al.*, 2016). Across the entire shortwave spectral region (i.e. 500-2400 nm), we also found reflectance displayed the highest variation (Figure 3b) in the visible region between about 500 and 700 nm (23%-32% across biomes), SWIR region between 1300 and 1700 nm (15%-17% across biomes), and far SWIR region between 1900 and 2500 nm (24%-37% across biomes). The NIR region displayed only minor variation of about 10% for all biome-level leaf spectra (Figure 3b).

Exploring the relationships between LMA and leaf spectral reflectance in the visible, NIR, and SWIR regions respectively offered insights into the strength and direction of the relationship in different parts of the spectrum (Figure S5). In the visible spectrum (Figure S5abc), the relationship between LMA and spectra is generally positive, as with the NIR (Figure S5d) despite the much lower coefficient of variation in that region (Figure 3b). There was a strong negative relationship between LMA and reflectance in the two SWIR (Figure S5ef) wavelengths (1800 & 2200 nm). Overall, these patterns are consistent with previous studies (le Maire *et al.*, 2008; Asner *et al.*, 2011; Feret *et al.*, 2011). For example, Feret et al., (2008) showed that spectral absorption of leaf dry matter content (LDMC, the analog used for LMA in the PROSPECT model) is generally

near zero in the visible and NIR but increases substantially when moving further out into the SWIR wavelengths. On the other hand, Slaton *et al.* (2001) illustrate the relationships between leaf morphology, thickness and other structural characteristics to changes in NIR reflectance across a range of leaf types and also found that there was a generally positive relationship between NIR reflectance (at 800 nm) and leaf structure parameters, including cuticle thickness and mesophyll properties. In addition, Serbin et al., (2014) showed the connection between select known and strong biochemical and structure absorption features in the 500-2400 nm wavelength range, and indicate how these relate to the prediction of traits, such as LMA.

Modeling multi-biome variation in LMA using leaf reflectance spectroscopy

Using the observed variation in measured LMA (Figure 2, Figure S1), leaf reflectance data (Figure 3a, Figure S4), and the covariation between the two (Figure S5) within our model development dataset we then evaluated the capacity to build a broad, multi-biome model to estimate LMA from leaf spectral reflectance across the broad range of plant species, growth environments, leaf ages, measurement approaches, and spectrometer instrumentation reflected in our data set. Our PLSR results (Figure 3cd, Figure 4) showed that a model based on leaf reflectance data was able to explain 89% of the variation in LMA in the independent validation data (and 91% in the calibration data using a cross-validation analysis, Figure S6a). Our multi-biome model also displayed a low overall independent validation error (RMSE=16 g m^{-2}) and bias (residual bias = 0.96 g m^{-2}, regression bias = 2.95 g m^{-2}), with the model calibration and validation residuals both centered around zero (Figure S6ef). Within each biome in the core calibration and validation datasets, the model fit also showed low error, ranging from a minimum RMSE of 12 g m^{-2} in the Arctic plants to 30 g m^{-2} in the Mediterranean plants (Figure S7). We further explored whether there were significant model biases attributable to variation with canopy position (Figure S8) and across leaf ages (Figure S9). Our results demonstrated that the model performance was very consistent across these different axes of variation, although, the leaf age evaluation

displayed some bias in the young (expanding leaves or current year foliage) and mature (one year old), and mature (> 1 year old) leaf age classes. Combined, the predictive error (dark black lines, 95% CI error bars on the points in Figure 4) for the estimated LMA is small, particularly for the most commonly observed values of LMA (e.g. LMA ≤ 150 g m^{-2}, Figure S1).

Our multi-biome LMA PLSR model has performance similar to, or better than previous studies using leaves from a much smaller range of species, locations, and growth environments (e.g. Asner *et al.*, 2011; Serbin *et al.*, 2014; Yang *et al.*, 2016). In addition, the model is able to cover a larger range of the global LMA trait-space (Figure 2; Figure S1), habitable climates (Figure 1) and variation in leaf optical properties (Figure S4) than earlier work. This includes previous studies utilizing semi-mechanistic leaf radiative transfer models (RTMs), including the PROSPECT model (e.g. Feret *et al.*, 2008; Feret *et al.*, 2011; Shiklomanov *et al.*, 2016). For example, Féret *et al.* (2018) found that performance of PROSPET-D LMA inversion varied based on the input dataset (and species within), inversion approach, and spectral domain, but that the overall PROSPECT-D inversion results across a smaller set of tropical to boreal samples were comparable to that shown here. Similarly, Shiklomanov et al., (2016) also found comparable results using a Bayesian inversion of PROSPECT-5b. However, in both cases, and with previous research (Malenovský *et al.*, 2006), thicker leaves and needles have hindered the inversion accuracy of semi-mechanist models like PROSPECT (e.g. Shiklomanov *et al.*, 2016) and as a result needle-leaf species are typically removed prior to analysis (e.g. Féret *et al.*, 2018). While approaches like PROSPECT are attractive to the remote sensing and plant trait ecology communities, issues surrounding the handling of more complex leaf morphologies and absorption properties need to be addressed in order to facilitate the confident use of these models global applications. On the other hand, our empirical approach was able to account for a broad range of morphologies and other drivers of leaf optical variation to produce a widely-applicable, multi-biome model (Figure 4).

The evaluation of the global model PLSR coefficient and VIP plots (Figure 3cd) highlight the regions of the spectrum that provide the strongest predictive power, many of which correspond to the areas showing the highest CV values, but note that the NIR region was also important. The SWIR spectral regions are known to contain absorption features related to structure, dry matter content, carbon compounds, and internal leaf water content (Curran, 1989) which also co-vary with LMA (Elvidge, 1990; Poorter *et al.*, 2009; de la Riva *et al.*, 2016). Our global model showed relatively high VIPs in the NIR region (Figure 3cd), consistent with previous studies that also highlighted the importance of this region (Asner *et al.*, 2011; Yendrek *et al.*, 2017), particularly in the transition between visible and NIR reflectance (~750-800 nm) and the region between 950-1200 nm. Corresponding PLSR coefficient values in the NIR were highest in the ~1000-1200 nm region. In general, the NIR has been shown to contain information connected to leaf internal scatter related to mesophyll layer thickness and water content, and varies with water, structural carbon, leaf thickness and variation in the epidermis layer (Ollinger, 2011), which also varies strongly with LMA (Castro-Díez *et al.*, 2000; Jacquemoud *et al.*, 2009; Poorter *et al.*, 2009; de la Riva *et al.*, 2016).

Our results show the powerful capability of the spectral approach to estimate a key leaf trait (LMA) across a high diversity of plant species, growth environments, leaf ages, leaf morphology, and biomes. Our synthesis represents the first time multiple datasets collected from different locations, by different groups, with different instrumentation across such large climatic and geographical ranges and from such a wide diversity of leaf types, including needle-leaf species (Figure 1), have been combined to test the capacity to generalize the spectral PLSR modeling approach. Importantly, there appears to be a general pattern in the PLSR models shown for this and previous studies where similar portions of the visible, NIR and SWIR regions display high importance in the estimation of LMA with spectra (Asner *et al.*, 2011; Serbin *et al.*, 2014; Yang *et al.*, 2016). This strongly suggests that the coordination between leaf optical properties

and traits can be used to develop generalized, global models for leaf traits using the spectroscopic approach.

External validation of the multi-biome LMA model

We further tested this capacity by applying our global leaf spectra-LMA model to three additional, completely independent datasets of measured leaf reflectance and LMA from trees, forbs, shrubs, and grasses, including the foundational LOPEX (Leaf Optical Properties Experiment) and ANGERS spectral datasets (Hosgood *et al.*, 1994; Jacquemoud *et al.*, 2003), which have been heavily used in remote sensing literature to develop and test the PROSPECT model (Feret *et al.*, 2008). We applied our model to the leaf reflectance and validated the predictions against the provided LMA data, respectively. For the LOPEX and ANGERS spectral datasets the results showed strong overall model performance (Figure S10) and low predictive error (RMSE of 21 g m^{-2}). Similarly, we applied our model to datasets collected in the upper Midwest and across NEON locations in the continental U.S., which further highlighted the high model performance and generality of the spectra-trait approach for estimating LMA across a range of species and environments (Figure S11). As such, using global models like the one presented here will not only avoid redundancy but also enable the capacity to collect much larger datasets by avoiding the need to directly, or destructively, sample traits. The spectral approach has the added benefit of providing the capacity to repeat measurements the same leaves during development or over a season, and during a manipulation or stress event because it does not require destructive harvesting.

Main applications of our results

Our study of multi-biome convergence in the leaf-level spectra-LMA relationships has four major implications. First, our framework and approach can largely be extended to other leaf traits. In addition to LMA, some other plant functional traits, e.g. pigments (chlorophyll a, b and

carotenoid), chemical concentration (water%, N%, C%, isotopic N and C), and carboxylation capacity (V_{cmax}), are also very important inputs for ESMs, and critical measures of plant form and function (Xu *et al.*, 2012; Rogers, 2014; van Bodegom *et al.*, 2014; Butler *et al.*, 2017; Rogers *et al.*, 2017a; Ricciuto *et al.*, 2018). Since these plant functional traits, have been shown to be connected to leaf spectral reflectance (Asner & Martin, 2008; Feret *et al.*, 2008; Kokaly *et al.*, 2009; Serbin *et al.*, 2012; Serbin *et al.*, 2014; Serbin *et al.*, 2015), we thus expect that similar globally convergent relationships with leaf spectra could be developed for these and other traits. To explore these additional generalities, future work can leverage emerging leaf spectra and trait databases, such as EcoSIS (https://ecosis.org), which will facilitate much faster and easier development, testing, and refinement of spectra-trait models.

Second, our work highlights the feasibility of using vegetation spectroscopy to advance large-scale monitoring of plant functional traits. This study evaluated the power of the spectroscopy approach at the leaf-level, but an important next step is to explore the ability to use a range of imaging spectroscopy platforms to map traits, such as LMA, at the canopy and landscape scales using airborne and spaceborne instruments across larger regions than currently explored, and ultimately globally (Schimel *et al.*, 2015). Here we show the capacity to generalize the spectral approach at the leaf scale, and there is strong evidence that similar approaches can be used at the canopy scale with imaging systems (Asner *et al.*, 2015; Singh *et al.*, 2015). This could augment or supplant the need for direct sampling and measurement of leaf traits at the global scale, especially given important upcoming satellite missions, e.g. Surface Biology and Geology Mission (National Academies of Sciences & Medicine, 2018), previously named *HyspIRI* (Lee *et al.*, 2015); EnMAP, (Guanter *et al.*, 2015). Furthermore, functional trait mapping from imaging spectroscopy could supplement methods using remote sensing combined with climatology (Butler *et al.*, 2017; Moreno-Martínez *et al.*, 2018)These capabilities would significantly enhance the use of trait observations to inform ESMs and would address critical needs in biodiversity monitoring (Skidmore *et al.*, 2015; Jetz *et al.*, 2016). However, lack of global

coverage, inconsistent processing workflows, and other challenges with the use of imaging spectroscopy has limited the ability to derive consistent global trait maps. Furthermore, approaches are needed to separate trait retrieval from climatology to allow for the characterization of biotic and abiotic drivers and allow for the mapping in the future under novel climates (Fisher *et al.*, 2015). This can only be realized with a spaceborne mapping imaging spectroscopy mission (National Academies of Sciences & Medicine, 2018)

Finally, we hypothesized that a single, multi-biome leaf reflectance model of LMA could be developed using datasets across diverse growth environments. We supported this hypothesis by showing a global convergence in the spectra-LMA relationship across samples representing a large portion of the global trait-space for LMA (Figures 4 and S1), which, when applied to external datasets from new, independent locations (Figures S10 & S11) showed similar model performance for estimating LMA. Importantly, since leaf traits, particularly LMA, display adaptions and acclimation to their growth environment (Poorter *et al.*, 2009; Osnas *et al.*, 2018), the success of spectra-LMA relationships, as shown in this study, further suggests that the spectroscopic approach could be an important, non-destructive means to help quantify and understand how plant traits acclimate to climatic variability and global change through rapid collection of traits, such as LMA, with spectra (Shiklomanov *et al.*, 2019).

Finally, the resulting global PLSR model (see supplementary R script and associated links within) developed here can be downloaded and used by the scientific community. Since measurement of leaf spectral reflectance is the only input required our model enables the rapid, accurate, and non-destructive estimation of LMA and can therefore be used for a broad range of additional applications e.g. monitoring plant response to an emerging stress or evaluating physiological traits of interest to breeders in high throughput phenotyping experiments (e.g. Silva-Perez *et al.*, 2017; Yendrek *et al.*, 2017; Ely *et al.*, 2019). As a result, databases such as EcoSIS can now be mined for increased coverage of important plant traits by applying this and other

spectra-trait models, potentially increasing the amount of available data for modeling and ecological research.

Acknowledgements

We thank Wil Lieberman-Cribbin & Jennifer Liebig for assistance with data collection in Barrow. We thank Neill Prohaska, Moira Hough, & Anthony John Junqueira Garnello for assistance with data collection in Biosphere 2. This work and associated field data collection campaigns were supported by the Next-Generation Ecosystem Experiments - NGEE Arctic, and NGEE Tropics - projects that are supported by the Office of Biological and Environmental Research in the Department of Energy, Office of Science, and through the United States Department of Energy contract No. DE-SC0012704 to Brookhaven National Laboratory, by NASA Earth and Space Sciences Fellowship (NNX08AV07H) provided to S.P.S., Forest Functional Types (NNX12AQ28G) and HyspIRI grants (NNX12AQ28G), NSF Macrosystems Biology grant (1638720) to P.A.T. and E.L.K., as well as a USDA McIntire-Stennis grant (WIS01809) to P.A.T. and E.L.K.

Author Contributions

S.P.S, J.W., and A.R. conceived the study, and defined the scope and focus of the manuscript. S.P.S developed the spectroscopic model of LMA. K.S.E., J.W., R.M., B.T.W, A.R, and S.P.S participated in the collection of data in the Arctic and tropical field sites (including the Biosphere 2 experiment). S.P.S, P.A.T, E.L.K, A.C. and Z.W. collected the datasets in California, upper Midwest, and at the NEON locations. SPS, JW and AR wrote the first draft of the manuscript. All authors contributed to the final version.

References

Asner GP, Martin RE. 2008. Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment* **112**(10): 3958-3970.

Asner GP, Martin RE, Anderson CB, Knapp DE. 2015. Quantifying forest canopy traits: Imaging spectroscopy versus field survey. *Remote Sensing of Environment* **158**(0): 15-27.

Asner GP, Martin RE, Tupayachi R, Emerson R, Martinez P, Sinca F, Powell GVN, Wright SJ, Lugo AE. 2011. Taxonomy and remote sensing of leaf mass per area (LMA) in humid tropical forests. *Ecological Applications* **21**(1): 85-98.

Brown J, Everett KR, Webber PJ, Maclean SF, Murray DF 1980. The Coastal Tundra at Barrow. In: Brown J, Miller PC, Tiezen LL, Bunnell FL eds. *An Arctic Ecosystem: the Coastal Tundra at Barrow, Alaska*. Stroudsburg, PA: Dowden, Hutchinson & Ross, Inc., 571.

Butler EE, Datta A, Flores-Moreno H, Chen M, Wythers KR, Fazayeli F, Banerjee A, Atkin OK, Kattge J, Amiaud B, et al. 2017. Mapping local and global variability in plant trait distributions. *Proceedings of the National Academy of Sciences* **114**(51): E10937-E10946.

Castro-Díez P, Puyravaud JP, Cornelissen JHC. 2000. Leaf structure and anatomy as related to leaf mass per area variation in seedlings of a wide range of woody plant species and types. *Oecologia* **124**(4): 476-486.

Chen S, Hong X, Harris CJ, Sharkey PM. 2004. Spare modeling using orthogonal forest regression with PRESS statistic and regularization. *IEEE Transaction on Systems, Man and Cybernetics* **34**: 898-911.

Curran PJ. 1989. Remote-sensing of foliar chemistry. *Remote Sensing of Environment* **30**(3): 271-278.

Dahlin KM, Asner GP, Field CB. 2013. Environmental and community controls on plant canopy chemistry in a Mediterranean-type ecosystem. *Proceedings of the National Academy of Sciences* **110**(17): 6895-6900.

de la Riva EG, Olmo M, Poorter H, Ubera JL, Villar R. 2016. Leaf Mass per Area (LMA) and Its Relationship with Leaf Structure and Anatomy in 34 Mediterranean Woody Species along a Water Availability Gradient. *Plos One* **11**(2): e0148788.

Dubois S, Desai AR, Singh A, Serbin SP, Goulden M, Baldocchi DD, Ma S, Oechel WC, Wharton S, Kruger EL, et al. 2018. Using imaging spectroscopy to detect variation in terrestrial ecosystem productivity across a water-stressed landscape. **in press**.

Elvidge CD. 1990. Visible and near-infrared reflectance characteristics of dry plant materials. *International Journal of Remote Sensing* **11**(10): 1775-1795.

Ely KS, Burnett AC, Lieberman-Cribbin W, Serbin S, Rogers A. 2019. Spectroscopy can predict key leaf traits associated with source–sink balance and carbon–nitrogen status.

Feret J-B, Francois C, Gitelson A, Asner GP, Barry KM, Panigada C, Richardson AD, Jacquemoud S. 2011. Optimizing spectral indices and chemometric analysis of leaf chemical properties using radiative transfer modeling. *Remote Sensing of Environment* **115**(10): 2742-2750.

Feret JB, Francois C, Asner GP, Gitelson AA, Martin RE, Bidel LPR, Ustin SL, le Maire G, Jacquemoud S. 2008. PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment* **112**(6): 3030-3043.

Féret JB, Gitelson AA, Noble SD, Jacquemoud S. 2017. PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. *Remote Sensing of Environment* **193**: 204-215.

Féret JB, le Maire G, Jay S, Berveiller D, Bendoula R, Hmimina G, Cheraiet A, Oliveira JC, Ponzoni FJ, Solanki T, et al. 2018. Estimating leaf mass per area and equivalent water

609 thickness based on leaf optical properties: Potential and limitations of physical modeling
610 and machine learning. *Remote Sensing of Environment*.
611 **Fisher JB, Huntzinger DN, Schwalm CR, Sitch S. 2014.** Modeling the Terrestrial Biosphere.
612 *Annual Review of Environment and Resources* **39**(1): 91-123.
613 **Fisher RA, Koven CD, Anderegg WRL, Christoffersen BO, Dietze MC, Farrior CE, Holm**
614 **JA, Hurtt GC, Knox RG, Lawrence PJ, et al. 2018.** Vegetation demographics in Earth
615 System Models: A review of progress and priorities. *Global Change Biology* **24**(1): 35-
616 54.
617 **Fisher RA, Muszala S, Verteinstein M, Lawrence P, Xu C, McDowell NG, Knox RG, Koven**
618 **C, Holm J, Rogers BM, et al. 2015.** Taking off the training wheels: the properties of a
619 dynamic vegetation model without climate envelopes, CLM4.5(ED). *Geosci. Model Dev.*
620 **8**(11): 3593-3619.
621 **Geladi P, Kowalski BR. 1986.** Partial least-squares regression - A tutorial. *Analytica Chimica*
622 *Acta* **185**: 1-17.
623 **Gitelson AA, Gritz Y, Merzlyak MN. 2003.** Relationships between leaf chlorophyll content and
624 spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher
625 plant leaves. *Journal of Plant Physiology* **160**(3): 271-282.
626 **Goulden ML, Anderson RG, Bales RC, Kelly AE, Meadows M, Winston GC. 2012.**
627 Evapotranspiration along an elevation gradient in California's Sierra Nevada. *Journal of*
628 *Geophysical Research-Biogeosciences* **117**.
629 **Guanter L, Kaufmann H, Segl K, Foerster S, Rogass C, Chabrillat S, Kuester T, Hollstein**
630 **A, Rossner G, Chlebek C, et al. 2015.** The EnMAP Spaceborne Imaging Spectroscopy
631 Mission for Earth Observation. *Remote Sensing* **7**(7): 8830.
632 **Hosgood B, Jacquemoud S, Andreoli G, Verdebout J, Pedrini G, Schmuck G. 1994.** Leaf
633 Optical Properties EXperiment 93 (LOPEX93). Ispra, Italy: European Commission —
634 Joint Research Centre.
635 **Jacquemoud S, Bidel L, Francois C, Pavan G 2003**. ANGERS Leaf Optical Properties
636 Database. http://opticleaf.ipgp.fr/index.php?page=database: OPTICLEAF.
637 **Jacquemoud S, Verhoef W, Baret F, Bacour C, Zarco-Tejada PJ, Asner GP, Francois C,**
638 **Ustin SL. 2009.** PROSPECT + SAIL models: A review of use for vegetation
639 characterization. *Remote Sensing of Environment* **113**(Suppl. 1, Sp. Iss. SI): S56-S66.
640 **Jetz W, Cavender-Bares J, Pavlick R, Schimel D, Davis FW, Asner GP, Guralnick R,**
641 **Kattge J, Latimer AM, Moorcroft P, et al. 2016.** Monitoring plant functional diversity
642 from space. *Nature Plants* **2**: 16024.
643 **John GP, Scoffoni C, Buckley TN, Villar R, Poorter H, Sack L, Maherali H. 2017.** The
644 anatomical and compositional basis of leaf mass per area. *Ecology Letters* **20**(4): 412-
645 425.
646 **Kattge J, DÍaz S, Lavorel S, Prentice IC, Leadley P, Bönisch G, Garnier E, Westoby M,**
647 **Reich PB, Wright IJ, et al. 2011.** TRY – a global database of plant traits. *Global*
648 *Change Biology* **17**(9): 2905-2935.
649 **Kokaly RF, Asner GP, Ollinger SV, Martin ME, Wessman CA. 2009.** Characterizing canopy
650 biochemistry from imaging spectroscopy and its application to ecosystem studies.
651 *Remote Sensing of Environment* **113**: S78-S91.
652 **le Maire G, François C, Soudani K, Berveiller D, Pontailler J-Y, Bréda N, Genet H, Davi H,**
653 **Dufrêne E. 2008.** Calibration and validation of hyperspectral indices for the estimation of
654 broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf
655 canopy biomass. *Remote Sensing of Environment* **112**(10): 3846-3864.
656 **Lebauer DS, Wang D, Richter KT, Davidson CC, Dietze MC. 2013.** Facilitating feedbacks
657 between field measurements and ecosystem models. *Ecological Monographs* **83**(2):
658 133-154.

Lee CM, Cable ML, Hook SJ, Green RO, Ustin SL, Mandl DJ, Middleton EM. 2015. An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities. *Remote Sensing of Environment* **167**: 6-19.

Liu M, Wang Z, Li S, Lü X, Wang X, Han X. 2017. Changes in specific leaf area of dominant plants in temperate grasslands along a 2500-km transect in northern China. *Scientific Reports* **7**(1): 10780.

Malenovský Z, Albrechtová J, Lhotáková Z, Zurita-Milla R, Clevers JGPW, Schaepman ME, Cudlín P. 2006. Applicability of the PROSPECT model for Norway spruce needles. *International Journal of Remote Sensing* **27**(24): 5315-5340.

Mevik B-H, Wehrens R. 2007. The pls Package: Principal Component and Partial Least Squares Regression in R. *Journal of Statistical Software* **18**(2): 1-24.

Moreno-Martínez Á, Camps-Valls G, Kattge J, Robinson N, Reichstein M, van Bodegom P, Kramer K, Cornelissen JHC, Reich P, Bahn M, et al. 2018. A methodology to derive global maps of leaf traits using remote sensing and climate data. *Remote Sensing of Environment* **218**: 69-88.

National Academies of Sciences E, Medicine. 2018. *Thriving on Our Changing Planet: A Decadal Strategy for Earth Observation from Space*. Washington, DC: The National Academies Press.

Niinemets U. 2007. Photosynthesis and resource distribution through plant canopies. *Plant Cell and Environment* **30**(9): 1052-1071.

Niinemets Ü, Keenan TF, Hallik L. 2015. A worldwide analysis of within-canopy variations in leaf structural, chemical and physiological traits across plant functional types. *New Phytologist* **205**(3): 973-993.

Ollinger SV. 2011. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist* **189**(2): 375-394.

Osnas JLD, Katabuchi M, Kitajima K, Wright SJ, Reich PB, Van Bael SA, Kraft NJB, Samaniego MJ, Pacala SW, Lichstein JW. 2018. Divergent drivers of leaf trait variation within species, among species, and among functional groups. *Proceedings of the National Academy of Sciences* **115**(21): 5480-5485.

Paula S, Pausas JG. 2006. Leaf traits and resprouting ability in the Mediterranean basin. *Functional Ecology* **20**(6): 941-947.

Pavlick R, Drewry DT, Bohn K, Reu B, Kleidon A. 2013. The Jena Diversity-Dynamic Global Vegetation Model (JeDi-DGVM): a diverse approach to representing terrestrial biogeography and biogeochemistry based on plant functional trade-offs. *Biogeosciences* **10**(6): 4137-4177.

Poorter H, Niinemets U, Poorter L, Wright IJ, Villar R. 2009. Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. *New Phytologist* **182**(3): 565-588.

R Core Team 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Reich PB, Ellsworth DS, Walters MB, Vose JM, Gresham C, Volin JC, Bowman WD. 1999. Generality of leaf trait relationships: A test across six biomes. *Ecology* **80**(6): 1955-1969.

Reich PB, Walters MB, Ellsworth DS. 1997. From tropics to tundra: Global convergence in plant functioning. *Proceedings of the National Academy of Sciences of the United States of America* **94**(25): 13730-13734.

Ricciuto D, Sargsyan K, Thornton P. 2018. The Impact of Parametric Uncertainties on Biogeochemistry in the E3SM Land Model. *Journal of Advances in Modeling Earth Systems* **10**(2): 297-319.

Rogers A. 2014. The use and misuse of V cmax in Earth System Models. *Photosynthesis Research* **119**(1-2): 15-29.

Rogers A, Medlyn BE, Dukes JS, Bonan G, von Caemmerer S, Dietze MC, Kattge J, Leakey ADB, Mercado LM, Niinemets Ü, et al. 2017a. A roadmap for improving the representation of photosynthesis in Earth system models. *New Phytologist* **213**(1): 22-42.

Rogers A, Serbin SP, Ely K 2016. Leaf Mass Area, Leaf Carbon and Nitrogen Content, Kougarok Road and Teller Road, Seward Peninsula, Alaska, 2016. *Next Generation Ecosystem Experiments Arctic Data Collection*. Oak Ridge National Laboratory: U.S. Department of Energy, Oak Ridge, Tennessee, USA.

Rogers A, Serbin SP, Ely KS, Sloan VL, Wullschleger SD. 2017b. Terrestrial biosphere models underestimate photosynthetic capacity and CO2 assimilation in the Arctic. *New Phytologist* **216**(4): 1090-1103.

Schimel D, Pavlick R, Fisher JB, Asner GP, Saatchi S, Townsend P, Miller C, Frankenberg C, Hibbard K, Cox P. 2015. Observing terrestrial ecosystems and the carbon cycle from space. *Global Change Biology* **21**(5): 1762-1776.

Serbin SP. 2012. *Spectroscopic determination of leaf nutritional, morphological, and metabolic traits.* Ph.D. Forestry Ph.D. Dissertation, University of Wisconsin - Madison Madison, Wisconsin.

Serbin SP, Dillaway DN, Kruger EL, Townsend PA. 2012. Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. *Journal of Experimental Botany* **63**(1): 489-502.

Serbin SP, Rogers A 2019. NGEE Arctic Leaf Spectral Reflectance, Kougarok Road, Seward Peninsula, Alaska, 2016. *Next Generation Ecosystem Experiments Arctic Data Collection*: U.S. Department of Energy, Oak Ridge, Tennessee, USA.

Serbin SP, Singh A, Desai AR, Dubois SG, Jablonski AD, Kingdon CC, Kruger EL, Townsend PA. 2015. Remotely estimating photosynthetic capacity, and its response to temperature, in vegetation canopies using imaging spectroscopy. *Remote Sensing of Environment* **167**: 78-87.

Serbin SP, Singh A, McNeil BE, Kingdon CC, Townsend PA. 2014. Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. *Ecological Applications* **24**(7): 1651-1669.

Shiklomanov A, Bradley B, Dahlin KM, Fox A, Gough CM, Hoffman FM, Middleton E, Serbin SP, Smallman L, Smith WK. 2019. Enhancing global change experiments through integration of remote-sensing techniques. *Frontiers in Ecology and the Environment* **in press**.

Shiklomanov AN, Dietze MC, Viskari T, Townsend PA, Serbin SP. 2016. Quantifying the influences of spectral resolution on uncertainty in leaf trait estimates through a Bayesian approach to RTM inversion. *Remote Sensing of Environment* **183**: 226-238.

Shipley B, Lechowicz MJ, Wright I, Reich PB. 2006. Fundamental trade-offs generating the worldwide leaf economics spectrum. *Ecology* **87**(3): 535-541.

Silva-Perez V, Molero G, Serbin SP, Condon AG, Reynolds MP, Furbank RT, Evans JR. 2017. Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. *Journal of Experimental Botany* **69**(3): 483-496.

Sims DA, Gamon JA. 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment* **81**(2): 337-354.

Singh A, Serbin SP, McNeil BE, Kingdon CC, Townsend PA. 2015. Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. *Ecological Applications* **25**(8): 2180-2197.

Skidmore AK, Pettorelli N, Coops NC, Geller GN, Hansen M, Lucas R, Mucher CA, O'Connor B, Paganini M, Pereira HM, et al. 2015. Agree on biodiversity metrics to track from space. *Nature* **523**(7561): 403-405.

760 **Slaton MR, Raymond Hunt Jr E, Smith WK. 2001.** Estimating near-infrared leaf reflectance
761 from leaf structural characteristics. *American Journal of Botany* **88**(2): 278-284.
762 **Townsend PA, Foster JR, Chastain RA, Currie WS. 2003.** Application of imaging
763 spectroscopy to mapping canopy nitrogen in the forests of the central Appalachian
764 Mountains using Hyperion and AVIRIS. *IEEE Transactions on Geoscience and Remote*
765 *Sensing* **41**(6): 1347-1354.
766 **Ustin SL, Roberts DA, Gamon JA, Asner GP, Green RO. 2004.** Using imaging spectroscopy
767 to study ecosystem processes and properties. *Bioscience* **54**(6): 523-534.
768 **van Bodegom PM, Douma JC, Verheijen LM. 2014.** A fully traits-based approach to modeling
769 global vegetation distribution. *Proceedings of the National Academy of Sciences*
770 **111**(38): 13733-13738.
771 **Volin JC, Kruger EL, Lindroth RL. 2002.** Responses of deciduous broadleaf trees to
772 defoliation in a CO2 enriched atmosphere. *Tree Physiology* **22**(7): 435-448.
773 **Walter A, Carmen Lambrecht S. 2004.** Biosphere 2 Center as a unique tool for environmental
774 studies. *Journal of Environmental Monitoring* **6**(4): 267-277.
775 **Whittaker RH. 1975.** *Communities and Ecosystems*: MacMillan Publishing Co., New York.
776 **Wold S, Sjostrom M, Eriksson L. 2001.** PLS-regression: a basic tool of chemometrics.
777 *Chemometrics and Intelligent Laboratory Systems* **58**(2): 109-130.
778 **Wright IJ, Reich PB, Westoby M, Ackerly DD, Baruch Z, Bongers F, Cavender-Bares J,**
779 **Chapin T, Cornelissen JHC, Diemer M, et al. 2004.** The worldwide leaf economics
780 spectrum. *Nature* **428**(6985): 821-827.
781 **Wright SJ, Horlyck V, Basset Y, Barrios H, Bethancourt A, Bohlman SA, Gilbert CS,**
782 **Goldstein G, Graham EA, Kitajima K. 2003.** Tropical canopy biology program,
783 Republic of Panama. Studying forest canopies from above: the International Canopy
784 Crane Network. Smithsonian Tropical Research Institute and UNEP.
785 **Wu J, Albert LP, Lopes AP, Restrepo-Coupe N, Hayek M, Wiedemann KT, Guan K, Stark**
786 **SC, Christoffersen B, Prohaska N, et al. 2016.** Leaf development and demography
787 explain photosynthetic seasonality in Amazon evergreen forests. *Science* **351**(6276):
788 972-976.
789 **Wu J, Chavana-Bryant C, Prohaska N, Serbin SP, Guan K, Albert LP, Yang X, Leeuwen**
790 **WJD, Garnello AJ, Martins G, et al. 2017.** Convergence in relationships between leaf
791 traits, spectra and age across diverse canopy environments and two contrasting tropical
792 forests. *New Phytologist* **214**(3): 1033-1048.
793 **Wullschleger SD, Epstein HE, Box EO, Euskirchen ES, Goswami S, Iversen CM, Kattge J,**
794 **Norby RJ, van Bodegom PM, Xu X. 2014.** Plant functional types in Earth system
795 models: past experiences and future directions for application of dynamic vegetation
796 models in high-latitude ecosystems. *Annals of Botany* **114**(1): 1-16.
797 **Xu C, Fisher R, Wullschleger SD, Wilson CJ, Cai M, McDowell NG. 2012.** Toward a
798 Mechanistic Modeling of Nitrogen Limitation on Vegetation Dynamics. *Plos One* **7**(5):
799 e37914.
800 **Xu X, Medvigy D, Powers Jennifer S, Becknell Justin M, Guan K. 2016.** Diversity in plant
801 hydraulic traits explains seasonal and inter-annual variations of vegetation dynamics in
802 seasonally dry tropical forests. *New Phytologist* **212**(1): 80-95.
803 **Yang X, Tang J, Mustard JF, Wu J, Zhao K, Serbin S, Lee J-E. 2016.** Seasonal variability of
804 multiple leaf traits captured by leaf spectroscopy at two temperate deciduous forests.
805 *Remote Sensing of Environment* **179**: 1-12.
806 **Yendrek CR, Tomaz T, Montes CM, Cao Y, Morse AM, Brown PJ, McIntyre LM, Leakey**
807 **ADB, Ainsworth EA. 2017.** High-Throughput Phenotyping of Maize Leaf Physiological
808 and Biochemical Traits Using Hyperspectral Reflectance. *Plant Physiology* **173**: 614-
809 626.
810

Figure Legends

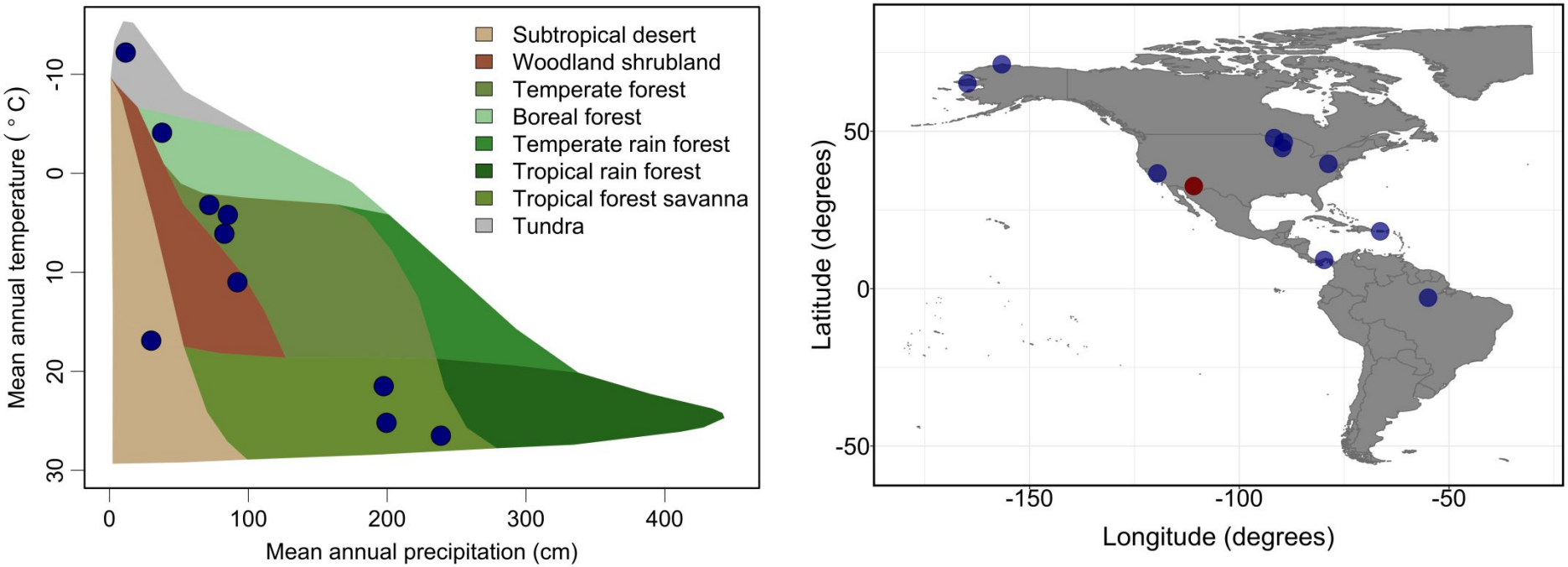
Figure 1 General location of the eleven sites where leaf reflectance and leaf mass area were measured in our model development dataset. Blue symbols show field sites, the red symbol shows the location of Biosphere 2 from which we sampled tropical species within the glasshouse environment. The left plot shows the sites plotted in climate space, binned by major climatic biomes (Whittaker, 1975) and the plot on the right displays the general location in geographic space.

Figure 2 Tukey box plots showing the range of leaf mass area (LMA, g m^{-2}) in our model development dataset calculated from measurements of leaf area and dry mass for our eleven regions, including the Biosphere2 glasshouse location. Sites are color coded by broad biome class (red, Arctic; green, boreal/temperate; purple, Mediterranean; orange, tropical). These are binned into four main biomes for clarity but correspond to the six Whittaker classes as shown in Figure 1. Box plots show the interquartile range (box), and median (solid horizontal line). The whiskers show lowest and highest datum still within 1.5 x inter quartile range of the lower and upper quartiles. Outliers are shown as black dots. Sample sizes by region: 609 Arctic leaves, 935 boreal/temperate leaves, 102 Mediterranean leaves (including 33 agricultural samples), and 832 tropical leaves (including 72 Biosphere2 samples)

Figure 3 Leaf reflectance and associated statistics. Panel A shows the mean leaf reflectance, 95% confidence interval (green shading) and minimum and maximum reflectance (dotted lines) in our model development dataset. Panel B shows the percent coefficient of variation for spectra from the four biomes represented in this study (red, Arctic; green, boreal/temperate; purple, Mediterranean; orange, tropical). Panel C shows the plot of the partial least squares regression (PLSR) model coefficients and Panel D shows the PLSR variable importance of prediction (VIP).

Figure 4 Observed leaf mass area (LMA) calculated from measured leaf area and dry mass versus LMA predicted using our spectral model. Our model development dataset (n=2,478) was split into two groups that were used to calibrate (black circles, n= 1978) and validate (grey circles, n=500) the model. The validation points are shown with ± 95% CI error bars. For clarity, validation points are shown layered on top of calibration points. The 1:1 line is shown as a broken black line and the predictive interval of the model is shown as the solid block lines. The regression between observed and predicted LMA is shown in blue (regression, thin blue line; 95% confidence interval, thick blue line). The R^2, root mean square error (RMSE) and regression bias (y axis intercept) for the validation data set are shown inside the panel. Data and statistics are presented in back transformed standard LMA units.

Figures



856
857

858 **Figure 1.**

859

860

861

862

863

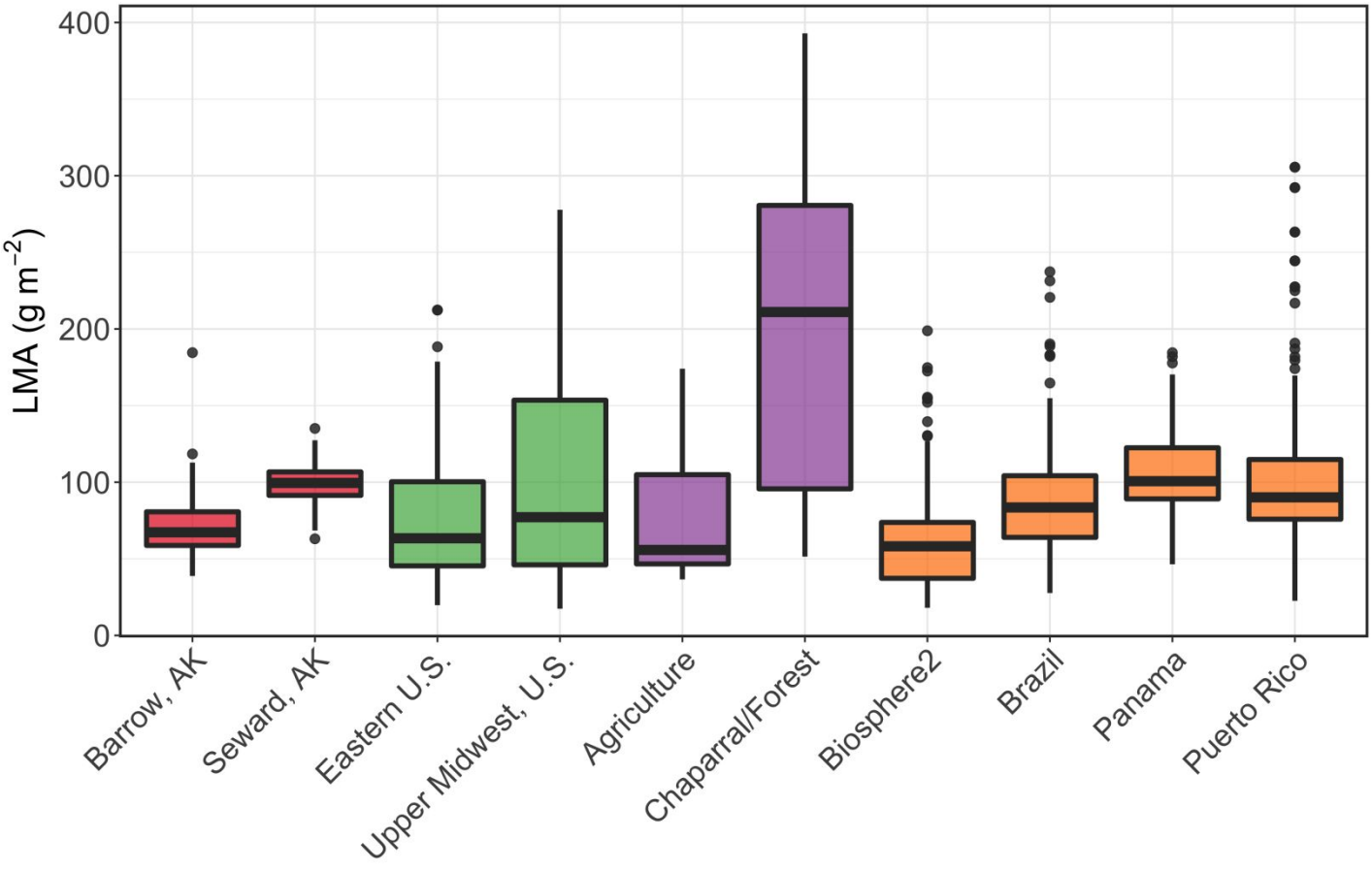


Figure 2.

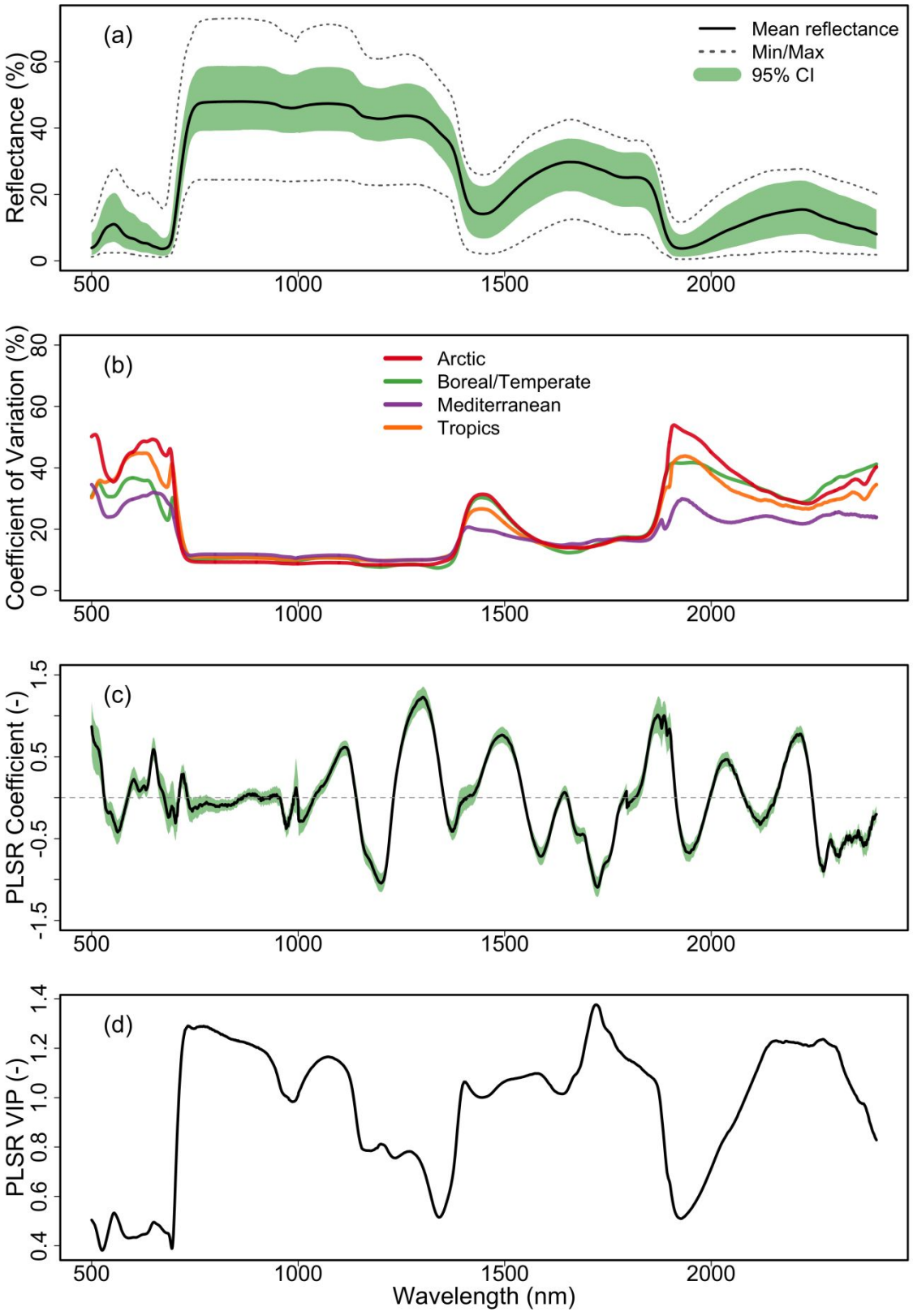
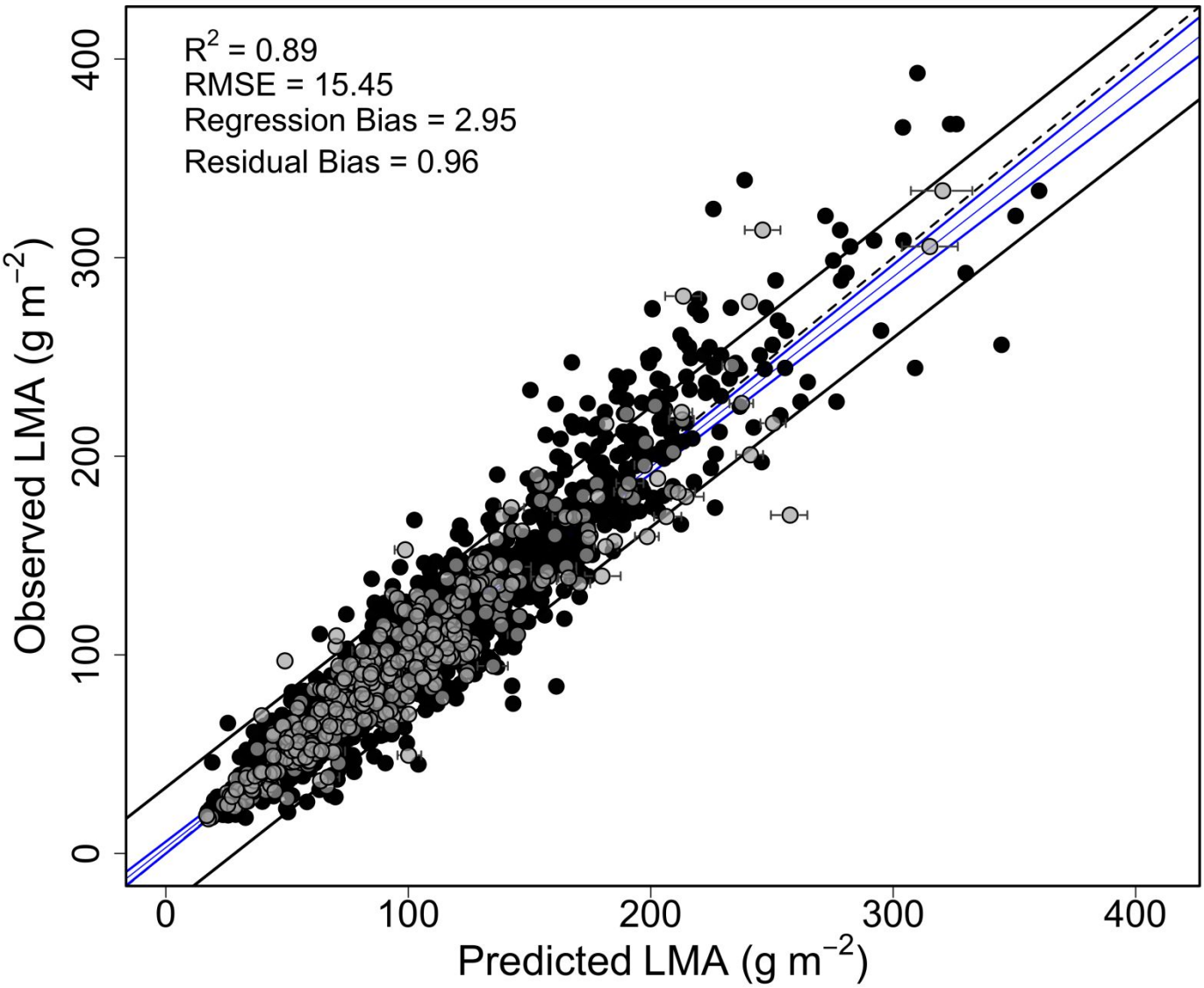

(a)
Mean reflectance
Min/Max
95% CI
Reflectance (%)
0
20
40
60
500
1000
1500
2000
(b)
Arctic
Boreal/Temperate
Mediterranean
Tropics
Coefficient of Variation (%)
0
20
40
60
80
500
1000
1500
2000
(c)
PLSR Coefficient (-)
-1.5
-0.5
0.5
1.5
500
1000
1500
2000
(d)
PLSR VIP (-)
0.4
0.6
0.8
1.0
1.2
1.4
500
1000
1500
2000
Wavelength (nm)

Figure 3.

876



877
878
879 **Figure 4.**